

【特許請求の範囲】

【請求項1】 入力された画像データおよび音声データを記録再生する半導体メモリを用いた画像および音声記録装置において、

上記画像データに対して圧縮符号化を施し、可変長の符号化画像データを生成する画像符号化手段と、

上記音声データに対して圧縮符号化を施し、可変長の符号化音声データを生成する音声符号化手段と、

着脱可能な半導体メモリに上記符号化画像データと上記符号化音声データとを記録する記録手段とからなることを特徴とする半導体メモリを用いた画像および音声記録装置。

【請求項2】 請求項1において、

上記半導体メモリに、上記符号化画像データと上記符号化音声データとを同期させるための時間情報も同時に記録することを特徴とする半導体メモリを用いた画像および音声記録装置。

【請求項3】 請求項1において、

上記半導体メモリに対して上記符号化画像データおよび上記符号化音声データをDOSファイル形式で記録するようにしたことを特徴とする半導体メモリを用いた画像および音声記録装置。

【請求項4】 入力された画像データおよび音声データとを記録再生する半導体メモリを用いた画像および音声記録再生装置において、

上記画像データに対して圧縮符号化を施し、可変長の符号化画像データを生成する画像符号化手段と、

上記音声データに対して圧縮符号化を施し、可変長の符号化音声データを生成する音声符号化手段と、

着脱可能な半導体メモリに上記符号化画像データと上記符号化音声データとを記録する記録手段と、

上記半導体メモリから上記符号化画像データを読み出し、上記読み出された符号化画像データを復号し、再生画像データを生成する画像復号手段と、

上記半導体メモリから上記符号化音声データを読み出し、上記読み出された符号化音声データを復号し、再生音声データを生成する音声復号手段とからなることを特徴とする半導体メモリを用いた画像および音声記録再生装置。

【請求項5】 請求項4において、

上記再生画像データは、ディスプレイモニタに表示され、上記再生音声データは、上記再生画像データと同期を取りスピーカから出力するようにしたことを特徴とする半導体メモリを用いた画像および音声記録再生装置。

【請求項6】 請求項4において、

上記半導体メモリに、上記符号化画像データと上記符号化音声データとを同期させるための時間情報も同時に記録することを特徴とする半導体メモリを用いた画像および音声記録装置。

【請求項7】 請求項4において、

上記半導体メモリに対して上記符号化画像データおよび上記符号化音声データをDOSファイル形式で記録するようにしたことを特徴とする半導体メモリを用いた画像および音声記録装置。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】この発明は、動画像および音声に不定長方式の圧縮を施す場合のデータの記録および再生を容易とすることが可能な半導体メモリを用いた画像および音声記録装置および記録再生装置に関する。

【0002】

【従来の技術】従来、動画像とその動画像に同期した音声とを記録する場合、動画像データと音声データを時分割（インターリーブ）で記録する手段が通常であった。これは、MPEG（Moving Picture Experts Group）1、2およびデジタルVCR（ビデオ・カセット・レコーダ）でみられる一般的な方法である。

【0003】MPEGでは、例えば図8Aに示すように、動画像データと音声データとを多重化する場合、パケットによる多重化を行う。例えば、動画像データを多重する場合、動画像データおよび音声データをそれぞれパケットと呼ばれる適当な長さのストリームに分割し、ヘッダ等の付加情報を付けて、適宜動画像データおよび音声データのバケットを切り換えて時分割伝送するようにされる。実際には、図8Bに示すように、動画像データおよび音声データは、複数のパケットを束ねたパックと呼ばれる構成単位で扱われる。その場合、パックヘッダの部分に同期再生用の時間基準参照用の付加情報等が付加される。

【0004】デジタルVCRでは、一例として図9に示すように、磁気テープ上に斜めのトラックとして記録される画像データは、DCT（離散的コサイン変換）による圧縮がされ、また外符号および内符号の2種類の誤り訂正符号を併用し、外符号を生成する前の1走査線分のシャフリングと、外符号を付加した後で内符号を生成する前の1セクタ内のシャフリングとが行われ、記録媒体に記録される。1トラックには、204ブロックの画像データが記録される。このブロックは、2ワードの同期符号、2ワードのID符号、85ワードの画像符号、8ワードの内符号パリティ、85ワードの画像符号および8ワードの内符号パリティから構成される。また、音声データは、48kHzで標準化された20ビットまでの符号を記録でき、動画像データとほぼ同じ構成となり、図に示すように動画像データの両端にギャップを介して、記録媒体に記録される。

【0005】

【発明が解決しようとする課題】このように、ディスクやテープに動画像および音声を記録する場合、光ディスク、HDD等に代表されるディスク媒体では、シーク時間が十分に速いとはいえず、またテープ媒体では、シー

クの概念すらなくらい、シーク時間が非常に長くなる問題があった。このような制約があり、再生信号の処理を容易とするためには、リアルタイムに、且つシーケンシャルに必然的にそれぞれのデータを固定長で多重化して記録する。

【0006】これに対して、ICメモ리카ードは、シーク概念自体なくらい、シーク時間が非常に高速なので、動画データおよび音声データを固定長データとして管理する必要はない。また、データ量を最小限にするには、データ量が少ないシーンないし無音状態も存在するので、圧縮法には、不定長方式を採用するのが合理的である。不定長方式を採用した場合、先頭からの連続再生のみであれば、動画と音声とを別ファイルとして管理する方法でも支障は生じないが、途中からの再生、途中までの巻戻しに相当することを行うとすれば、何らかの時間情報が必要である。

【0007】従って、この発明の目的は、記録媒体としてICメモ리카ード等の半導体メモリを使用して動画と音声との同期をとることが可能な半導体メモリを用いた画像および音声記録装置および記録再生装置を提供することにある。

【0008】

【課題を解決するための手段】請求項1に記載の発明は、入力された画像データおよび音声データを記録再生する半導体メモリを用いた画像および音声記録装置において、画像データに対して圧縮符号化を施し、可変長の符号化画像データを生成する画像符号化手段と、音声データに対して圧縮符号化を施し、可変長の符号化音声データを生成する音声符号化手段と、着脱可能な半導体メモリに符号化画像データと符号化音声データとを記録する記録手段とからなることを特徴とする半導体メモリを用いた画像および音声記録装置である。

【0009】請求項4に記載の発明は、入力された画像データおよび音声データとを記録再生する半導体メモリを用いた画像および音声記録再生装置において、画像データに対して圧縮符号化を施し、可変長の符号化画像データを生成する画像符号化手段と、音声データに対して圧縮符号化を施し、可変長の符号化音声データを生成する音声符号化手段と、着脱可能な半導体メモリに符号化画像データと符号化音声データとを記録する記録手段と、半導体メモリから符号化画像データを読み出し、読み出された符号化画像データを復号し、再生画像データを生成する画像復号手段と、半導体メモリから符号化音声データを読み出し、読み出された符号化音声データを復号し、再生音声データを生成する音声復号手段とからなることを特徴とする半導体メモリを用いた画像および音声記録再生装置である。

【0010】動画データが、例えばMPEGで符号化され、GOP単位でICメモ리카ードに記憶され、そのGOP単位に同期して音声データが、例えばATRAC

で符号化され、ICメモ리카ードに記憶され、さらに圧縮画像データと圧縮音声データとの同期をとるための時間情報管理が記憶される。これによって、任意の位置の再生を行うときに、時間情報管理から圧縮画像データおよびその圧縮画像データに同期した圧縮音声データを任意に読み出し、再生することができる。

【0011】

【発明の実施の形態】以下、この発明の一実施例を図面を参照して詳細に説明する。この発明の一実施例の構成を図1に示す。1で示すマイクから音声が入力される。入力された音声は、マイク1に含まれる音声/電気信号変換素子によって電気信号に変換され、音声信号として増幅器2へ供給される。増幅器2では、音声信号が増幅され、必要な帯域に制限される。その後、音声信号は、A/D変換回路3で必要な帯域の2倍以上のサンプリング周波数でデジタル化される。この音声データは、音声データ圧縮/伸長回路4の音声圧縮部で符号化され、DRAM等からなるメモリ14に記録される。

【0012】メモリ14に記録された圧縮音声データは、カードI/F回路15を介して読み出され、ICメモ리카ード16で記録される。このICメモ리카ード16は、一般的なメモ리카ードの機能を備え、且つパーソナル・コンピュータで読み書き可能な記憶媒体の1つである。また、一例として、ICメモ리카ード16は、DOS形式のフォーマットで記録/再生がなされる。実際には、カードI/F回路15とICメモ리카ード16の間にICメモ리카ード16を装着するコネクタが配され、コネクタによってICメモ리카ードが着脱自在とされる。

【0013】ICメモ리카ード16に記録された圧縮音声データは、カードI/F回路15を介して読み出される。読み出された圧縮音声データは、メモリ14の音声用の記録領域に書き込まれる。書き込まれた圧縮音声データは、音声データ圧縮/伸長回路4の音声伸長部において、リアルタイムに伸長される。伸長された音声データは、D/A変換器5へ供給され、アナログ化される。アナログ化された音声信号は、増幅器6で増幅され、オーディオ信号としてスピーカ7から出力される。

【0014】また、動画となる被写体は、レンズ、絞り部8から入力される。被写体は、レンズ、絞り部8からCCD撮像素子9へ入力される。CCD撮像素子9では、被写体が電荷として蓄積され、電気信号に変換される。電気信号は、動画信号としてCDS回路10へ供給される。CDS回路10は、相関二重サンプリング回路であり、ノイズの低減を図るものである。また、CPU22は、レンズ、絞り部8の機械系のドライバ18、CCD撮像素子9、CDS回路10およびA/D変換器11の電気系のタイミング信号発生部(TG)19を制御する。

【0015】CDS回路10の出力信号は、A/D変換

器11に供給され、A/D変換器11では、最適なサンプリング周波数、例えばNTSC信号のサブキャリア周波数の整数倍となるようなサンプリング周波数にてデジタル化される。デジタル化されたCDS回路10の出力信号は、デジタル信号処理回路12へ供給される。デジタル信号処理回路12では、供給された信号に対してガンマ補正、色分離等の通常のカメラ信号処理がなされ、輝度信号Y、色差信号Cb、Crが作成される。これらの信号は、画像データ圧縮/伸長回路13の画像圧縮部で符号化され、記憶素子であるメモリ14に記録される。メモリ14に記録された圧縮画像データは、カードI/F回路15を介して読み出され、ICメモリカード16へ出力され、記録される。

【0016】ICメモリカード16に記録された圧縮画像データは、カードI/F回路15を介して読み出される。読み出された圧縮画像データは、メモリ14の画像用の記録領域に書き込まれる。書き込まれた圧縮画像データは、画像データ圧縮/伸長回路13の画像伸長部において、リアルタイムに伸長される。伸長された動画データは、デジタル信号処理回路12へ供給される。デジタル信号処理回路12へ供給された輝度信号Y、色差信号Cb、Crは、NTSC信号にデジタル・デコードされ、D/A変換されたビデオ信号としてLCD（液晶ディスプレイ）17へ出力される。

【0017】操作部21からの指示、またはリモコン（図示せず）等の外部動作指示に従い、CPU22のカメラ内部の動作を制御する。また、カメラ内部状態等の表示は、例えばLCD、LED、EL等から構成される表示部20に表示される。

【0018】ICメモリカード16のメモリ領域の一例を図2に示す。このICメモリカードは、ブート領域M1、メモリアロケーションテーブル領域M2、ディレクトリ情報テーブル領域M3、ファイル情報テーブルM4およびファイルデータ領域M5から構成される。ブート領域M1には、ブートブロック情報が記録される。このブートブロック情報には、このICメモリカードのバージョン、タイトル、パスワード、書き換え日付、書き換え機種、パーティション情報、ルートディレクトリ、メモリアロケーションのポインタおよびサイズ、ディレクトリ情報のポインタおよびサイズ、ファイル情報のポインタおよびサイズ等が記録される。

【0019】メモリアロケーションテーブル領域M2には、メモリブロック情報が記録される。このメモリブロック情報は、このICメモリカードの記録領域を所定のサイズの複数のブロックに分割した、各ブロックの情報が記録される。例えば、ユーザにより記録されたファイルデータが複数のブロックに跨がって記録されるような場合、このブロックが先頭ブロックなのか先頭ブロックからリンクされたブロックなのかを示す情報、このブロックが使用されているのか否かの情報、このブロックが

予め設定された消去回数より多く消去されている場合、データの信頼性を考えて使用不可とする情報が記録される。

【0020】また、このブロックの消去回数、このブロックのタイプ、例えばブート領域に使用されているブロックなのか、メモリ領域なのか、ディレクトリ情報テーブルM3に使用されているブロックなのか、ファイル情報テーブルM4に使用されているブロックなのか、ユーザによってファイルデータが記録されているブロックなのか、が情報として記録される。上述したように、リンクされているブロックがある場合、次のブロックのアドレスが情報として記録され、このブロックで終了している場合、終了という情報が記録される。

【0021】ディレクトリ情報テーブル領域M3には、メモリブロック情報でディレクトリ情報テーブルに使用されているディレクトリとされたディレクトリ情報が記録される。例えば、ユーザにより記録されたファイルデータが複数のディレクトリに跨がって記録されるような場合、このディレクトリが先頭なのか先頭のディレクトリからリンクされたディレクトリなのかを示す情報、このディレクトリが使用されているのか否かの情報、このディレクトリ内に記録されるファイルの数を予め設定した場合、既に記録されているファイルの情報を要素0（ファイル番号）、要素1（ファイル番号）、・・・、要素n（ファイル番号）として記録し、上述したように、リンクされているディレクトリがある場合、次のディレクトリのアドレスが情報として記録され、このディレクトリで終了している場合、終了という情報が記録される。

【0022】ファイル情報テーブル領域M4には、メモリブロック情報でファイル情報テーブルに使用されているファイルとされたファイル情報が記録される。例えば、このファイルがファイルとして使用されているのか、ディレクトリとして使用されているのか、未使用なのか記録され、オーナーID、このファイルが作成された作成日または修正された修正日、このファイルを読み出し専用ファイルにするか、上書き可能なファイルにするか、またはプログラムの場合、実行することができないようにするかが記録される。さらに、このファイルのサイズ、名前、メモリブロックの番号またはディレクトリの番号、またファイルが記録されている場合、ファイルの属性およびクリエイター情報などが記録される。このクリエイター情報とは、例えばどのようなアプリケーションでこのファイルが作成されたか、またはどのようなアプリケーションでこのファイルが使用可能かなどが記録される。

【0023】ファイルデータM5には、メモリブロック情報でユーザによってファイルデータに使用されているファイルデータが記録される。このファイルデータは、圧縮画像データおよび圧縮音声データ等が記録される。

これらの管理方法の一例として、ICメモ리카ード16の絶対アドレスを使用する。

【0024】この一実施例では、画像データ圧縮／伸長回路13では、動画データ圧縮のためにMPEG2の符号化／復号を使用し、音声データ圧縮／伸長回路4では、音声データを圧縮するためにATRACの符号化／復号を使用する。符号化した動画データおよび音声データは、上述したICメモ리카ード16のファイルデータM5に記録される。そのデータフォーマットの一例を説明する。ICメモ리카ード16は、パーソナル・コンピュータで記録／再生を可能とするために、パーソナル・コンピュータ側のオペレーション・システムで記録／再生し易いフォーマットで記録される。一般的には、DOSフォーマットが使用される。よって、この実施例でも、DOSファイル形式を適用する。

【0025】ここで、簡単にATRACを説明する。アナログの音声信号を44.1kHzのサンプリング周波数でサンプリングし、デジタル化された音声データを最大11.6msの時間窓で切り出し、変形DCT演算で周波数成分に分解する。変形DCT演算で生じやすいプリエコーを避けるため周波数帯域を3つに分ける。変形DCTで周波数軸に変換した後に、人間の聴覚特性を利用してデータの間引きを行う。聴覚マスキングを利用すると大きい音楽信号周波数成分の近くの雑音や小さい信号は聞こえないので聞こえる成分は残し、聞こえない必要でない成分はカットする。そして、圧縮したデータが1サウンドグループ(SG)毎に束ねてクラスタ単位で記録媒体に記録される。

【0026】このATRACは、MD(ミニディスク)で採用され、標準化されている音声データの圧縮の方法である。このMDでは、CD-ROMのモード2が利用され、録音用光磁気ディスクに圧縮音声データが記録される。録音用光磁気ディスクに記録されるMDのデータ構成は、図3に示す。この録音用光磁気ディスクでは、図3Aに示すように、36セクタがひと固まり(=1クラスタ)として記録される。先頭の3セクタはリンク・セクタ領域とされ、4セクタ目はサブデータ・セクタ領域とされ、残りの32セクタは圧縮データ用の領域として扱われる。また、図3Bに示すように、1セクタは、2352バイトからなり、その内データ用は、2332バイトである。圧縮音声データは、424バイトを1単位として扱い、その単位をサウンドグループと呼ぶ。図3Cに示すように、11個のサウンドグループが2セクタに記録される。そして、図3Dに示すように、1サウンドグループの圧縮を解き、時間軸情報に戻すと左右チャンネルの512サンプルとなる。

【0027】図4に示すように、動画データにMPEG2を施し、音声データにATRACを施した場合であり、動画データと音声データとを別のファイルとし、動画データと音声データとの同期をとるための時

間情報管理ファイルの3つのファイルが1組のデータとなる。図4Aに示す、動画データファイルは、ヘッダ領域と動画データの領域とから構成される。画像ファイルのヘッダ領域には、一例として、記録されている動画データの記録時間、画面サイズ、フレーム数、記録ビットストリームIDおよび画像構成等が記録されている。画像構成とは、輝度信号Yおよび色差信号Cb、Crのサンプリング周波数の比であって、例えば4:2:2または4:1:1となる。動画データは、この一実施例では、MPEGのGOP単位に分けられる。

【0028】図4Bに示す、音声ファイルは、ヘッダ領域と音声データの領域とから構成される。音声ファイルのヘッダ領域には、一例として、記録されている音声データの記録時間、サンプリング周波数、量子化ビット数および記録ビットストリームID等が記録されている。音声データは、上述したATRACのクラスタを集めたオーディオグループ(AG)単位に分けられる。

【0029】そして、図4Cに示す、時間情報管理ファイルも同様に、ヘッダ領域と音声データの領域とから構成される。時間情報管理ファイルのヘッダ領域には、一例として、記憶されている時間情報管理ファイルのデータサイズおよびデータ種別ID等が記録されている。時間情報データは、時間管理情報であり、固定サイズのブロックで構成され、例えばGOP単位での動画データに同期した音声データのデータアドレスを下記に示す内容で記録される。

【0030】GOP-Nのスタートアドレスが記録され、GOP-Nと同期しているAG番号(M)が記録される。そして、AG-Mのスタートアドレスが記録され、そのAG-Mの中のオフセットが記録される。

【0031】具体的に、動画データに圧縮を施すMPEG2のGOPを基準にして、音声データがオーディオグループに分割される。GOP、オーディオグループおよび時間情報ファイルの一例を図5に示す。図5Aに示すように、動画データファイルの場合、動画データにMPEG2の圧縮を施し、GOP単位で記録される。また、不定長圧縮法のMPEG2を用いているため、各GOPに対応するデータの記録領域の長さが異なる。このように、動画データは、それぞれのデータサイズが異なる複数のフレーム単位(GOP単位)で記録される。

【0032】図5Bに示すように、音声ファイルの場合、音声データにATRACの圧縮を施し、図3Bに示すような、データ構成とされる。例えば、1GOPに対応する圧縮されたオーディオグループは、不定長である。このように、音声データもGOPと対応するサンプルの集合(オーディオグループ)で記録され、このオーディオグループのデータサイズも異なる。

【0033】そして、図5Cに示すように、時間情報管理ファイルの場合、動画データのGOPと音声データのオーディオグループとの同期をとるための時間情報が

記録される。この時間情報データもGOP単位にデータが分割される。この実施例では、GOP番号(n)とオーディオグループ番号(m)は、時間情報番号(p)に時間情報が記録され、時間情報番号毎に動画像データとその動画像データに同期した音声データの開始アドレスが格納されている。また、1GOPと付随するオーディオデータが1オーディオグループの場合には、GOP番号(n)とオーディオグループ番号(m)と同一となるが、1GOPが複数のオーディオグループと対応しても良い。

【0034】このように、圧縮画像データと圧縮音声データとが異なるファイルで記録されたICメモリカードから動画像および音声再生する一例を説明する。図6は、ICメモリカードの再生専用機の一例を示したブロック図である。ICメモリカード32は、まずカードI/F回路33によって再生される時間情報管理ファイルが読み出される。読み出された時間情報管理ファイルに従って、圧縮画像データおよび圧縮音声データが読み出される。圧縮音声データは、音声データ圧縮/伸長回路34へ供給され、圧縮画像データは、画像データ圧縮/伸長回路38へ供給される。

【0035】画像データ圧縮/伸長回路36では、圧縮画像データが伸長され、動画像データとしてデジタル信号処理回路39へ供給される。デジタル信号処理回路39では、供給された動画像データがアナログ化されたNTSC方式の信号へ変換され、LCD40へ供給される。LCD40では、動画像信号が表示される。音声データ圧縮/伸長回路34では、圧縮音声データが伸長され、音声データとしてD/A変換器35へ供給される。アナログ化された音声信号は、増幅器36を介してスピーカ37からLCD40に表示される動画像信号と同期をとり出力される。これらの制御は、CPU31によって行われる。

【0036】この再生時のCPUの制御の一例を図7のフローチャートを用いて説明する。ステップS1では、ICメモリカードから時間情報管理ファイルが読み出され、CPU22のメモリに記録される。再生を希望する位置が入力されると、ステップ2では、メモリに記録された時間情報管理ファイルから再生を希望する圧縮画像データのスタートアドレス(スタートポイント)が算出される。ステップ3では、同様に記録されている時間情報管理ファイルから再生する圧縮音声データの開始アドレス(スタートポイント)が算出される。

【0037】ステップS4では、算出された圧縮画像データの開始アドレスから圧縮画像データが読み出される。読み出された圧縮画像データが伸長処理部へ送ら

る。ステップS5では、算出された圧縮音声データの開始アドレスから圧縮音声データが読み出される。読み出された圧縮音声データが伸長処理部へ送られる。このとき、音声データにオフセットがある場合、音声データは、オフセット分ずらして伸長される。これによって、再生された動画像データと音声データの同期は、GOP単位で取られている。

【0038】この一実施例では、動画像データを基準として同期を行ったが、音声データを基準として同期を行っても良い。

【0039】また、この一実施例では、いわゆるDOSファイルによる管理を前提に説明したが、それ以外の管理方式による管理でも同様に実現できる。

【0040】

【発明の効果】この発明に依れば、任意の時間の動画像と音声を再生する場合、その時間の時間情報ファイルを読み出し、その時間情報ファイルに記録された動画像データと音声データのスタートアドレスに応じて動画像データと音声データを同期させて再生することができる。

【図面の簡単な説明】

【図1】この発明の一実施例のブロック図である。

【図2】この発明に係るICメモリカードのメモリの概略図である。

【図3】ATRACの説明に用いる略線図である。

【図4】この発明に係るICメモリカードに記録されるデータの説明に用いる略線図である。

【図5】この発明に係るICメモリカードに記録されるデータの説明に用いる略線図である。

【図6】この発明の再生時の一例を説明するための概略的なブロック図である。

【図7】この発明の再生時の一例を説明するためのフローチャートである。

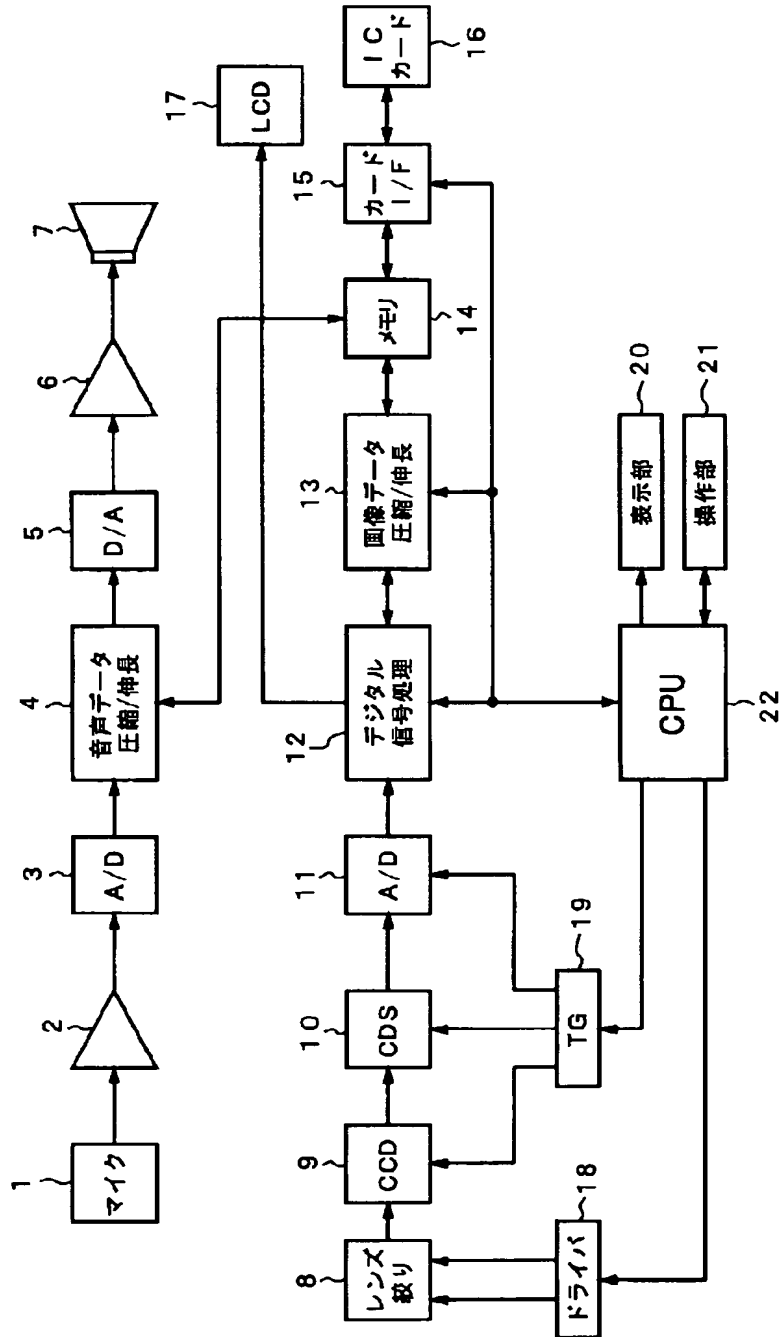
【図8】MPEGの説明に用いる略線図である。

【図9】デジタルVCRの説明に用いる略線図である。

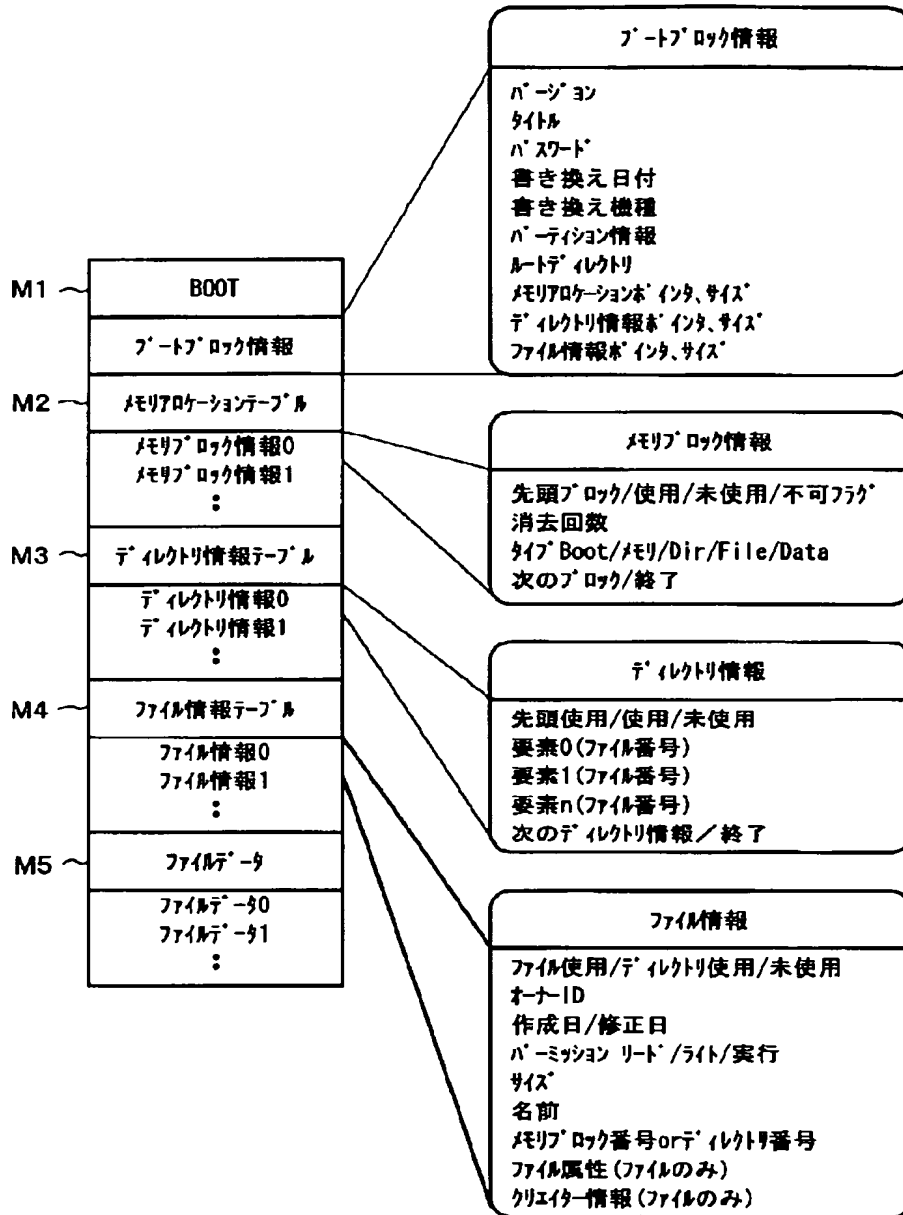
【符号の説明】

1・・・マイク、2、6・・・増幅器、3、11・・・A/D変換器、4・・・音声データ圧縮/伸長回路、5・・・D/A変換器、7・・・スピーカ、8・・・レンズ、絞り部、9・・・CCD撮像素子、10・・・CDS回路、12・・・デジタル信号処理回路、13・・・画像データ圧縮/伸長回路、14・・・メモリ回路、15・・・カードI/F回路、16・・・ICメモリカード、17・・・LCD、18・・・ドライバ、19・・・タイミング信号発生部、20・・・表示部、21・・・操作部、22・・・CPU

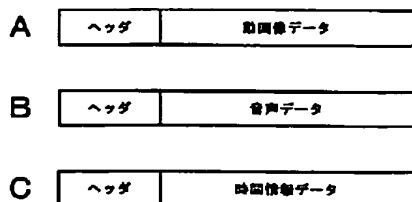
【図1】



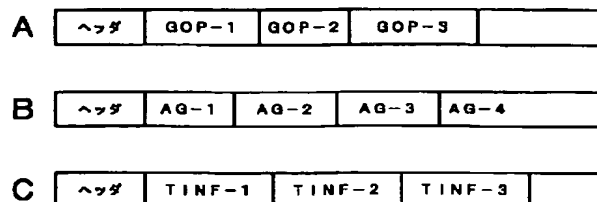
【図2】



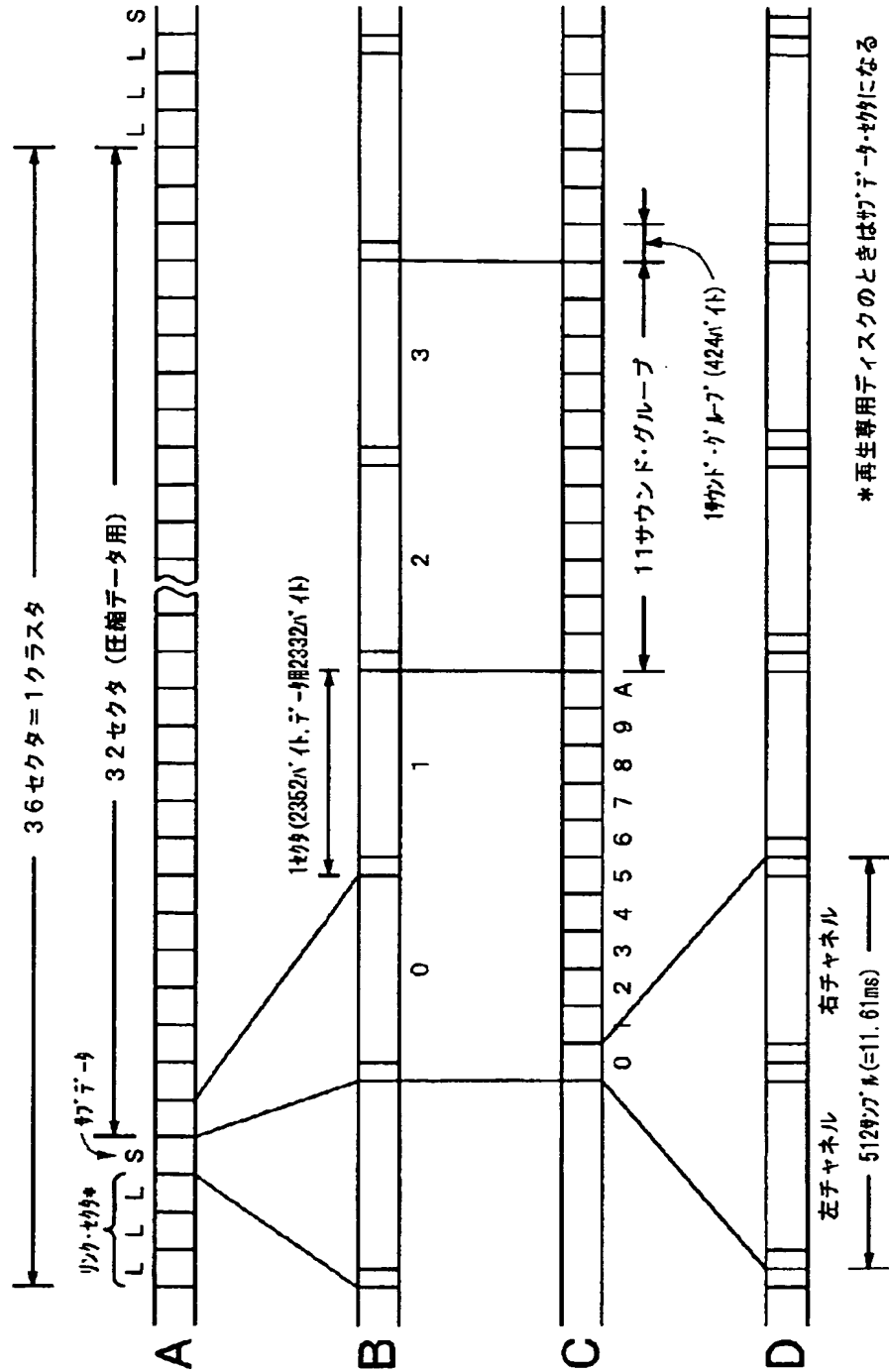
【図4】



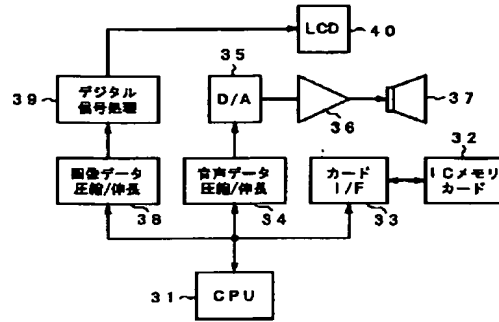
【図5】



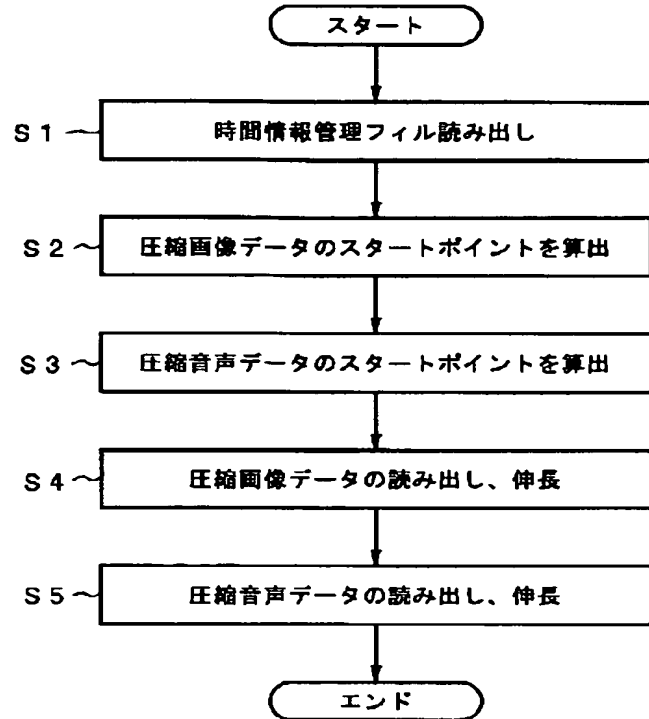
【図3】



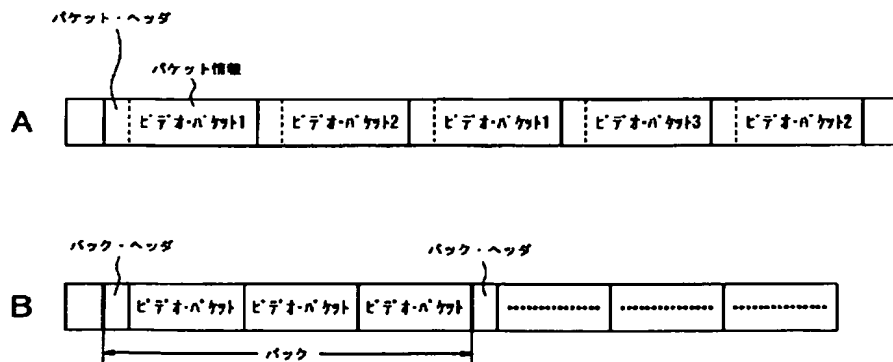
【図6】



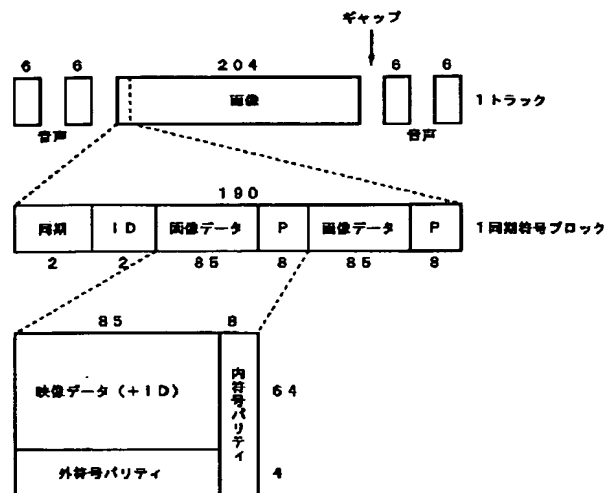
【図7】



【図8】



【図9】



PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-210409

(43)Date of publication of application : 07.08.1998

(51)Int.Cl. H04N 5/92
G10L 9/18
H04N 5/907

(21)Application number : 09-005844 (71)Applicant : SONY CORP

(22)Date of filing : 16.01.1997 (72)Inventor : KOBAYASHI TOSHIHARU

(54) IMAGE AND VOICE RECORDER AND IMAGE AND VOICE RECORDER AND REPRODUCER USING SEMICONDUCTOR MEMORY

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a recorder and a recorder/reproducer which can secure synchronization between dynamic images and voices by recording the encoded and compressed image and voice data on an IC memory card together with the time information that secures the synchronization between those image and voice data.

SOLUTION: Voice data are encoded at a voice compression part of a voice data compression/expansion circuit 4 and recorded in a memory 14 consisting of a DRAM etc. Then the compressed voice data are read out the memory 14 via a card I/F circuit 15 and recorded on an IC memory card 16. The luminance and color difference signals produced from the dynamic image signals are encoded at an image compression part of an image compression/expansion circuit 13 and recorded in the memory 14. The compressed image data are read out of the memory 14 via the circuit 15 and recorded on the card 16. Furthermore the time management information is recorded on the card 16 to secure synchronization between the compressed image and the voice data.

CLAIMS

[Claim(s)]

[Claim 1] A picture and a voice recording device characterized by comprising the following using semiconductor memory which carries out record reproduction of image data and voice data which were inputted.

An image encoding means which performs compression encoding to described image data and generates variable-length coded image data.

A voice coding means to perform compression encoding to the above-mentioned

voice data and to generate variable-length coded voice data.

A recording device which records the above-mentioned coded image data and the above-mentioned coded voice data on removable semiconductor memory.

[Claim 2] A picture and a voice recording device using semiconductor memory also recording simultaneously a hour entry for synchronizing the above-mentioned coded image data and the above-mentioned coded voice data with the above-mentioned semiconductor memory in claim 1.

[Claim 3] A picture and a voice recording device using semiconductor memory recording the above-mentioned coded image data and the above-mentioned coded voice data in DOS file form to the above-mentioned semiconductor memory in claim 1.

[Claim 4] A picture and voice recording playback equipment characterized by comprising the following using semiconductor memory which carries out record reproduction of image data and voice data which were inputted.

An image encoding means which performs compression encoding to described image data and generates variable-length coded image data.

A voice coding means to perform compression encoding to the above-mentioned voice data and to generate variable-length coded voice data.

A recording device which records the above-mentioned coded image data and the above-mentioned coded voice data on removable semiconductor memory.

An image decoding means which reads the above-mentioned coded image data from the above-mentioned semiconductor memory decodes coded image data by which reading appearance was carried out [above-mentioned] and generates reproduced image data. A speech decoding means which reads the above-mentioned coded voice data from the above-mentioned semiconductor memory decodes coded voice data by which reading appearance was carried out [above-mentioned] and generates playback voice data.

[Claim 5] A picture and voice recording playback equipment using semiconductor memory characterized by having displayed the above-mentioned reproduced image data on a display monitor and the above-mentioned playback voice data's taking the above-mentioned reproduced image data and a synchronization and making it output them from a loudspeaker in claim 4.

[Claim 6] A picture and a voice recording device using semiconductor memory also recording simultaneously a hour entry for synchronizing the above-mentioned coded image data and the above-mentioned coded voice data with the above-mentioned semiconductor memory in claim 4.

[Claim 7] A picture and a voice recording device using semiconductor memory recording the above-mentioned coded image data and the above-mentioned coded voice data in DOS file form to the above-mentioned semiconductor memory in claim 4.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the picture/voice recording device and recording and reproducing device using the semiconductor memory which can make easy record and reproduction of the data in the case of compressing an unfixed length method into video and a sound.

[0002]

[Description of the Prior Art] When recording the sound in sync with video and its video conventionally, a means to record dynamic image data and voice data by time sharing (interleave) was usual. This is a general method seen with MPEG (Moving Picture Experts Group) 1 and 2 and digital VCR (video cassette recorder).

[0003] In MPEG as shown for example in drawing 8 A when multiplexing dynamic image data and voice data, multiplexing by a packet is performed. For example, when carrying out multiplex [of the dynamic image data], dynamic image data and voice data are divided into the stream of the suitable length called a packet, respectively, additional informations such as a header is attached to the packet of dynamic image data and voice data is switched suitably, and time division transmission is made to be carried out. Actually, as shown in drawing 8 B, dynamic image data and voice data are treated by the constitutional unit called the pack which bundled two or more packets. In that case, the additional information for referring to the time base for synchronous reproduction etc. are added to the portion of a pack header.

[0004] In digital VCR as shown in drawing 9 as an example, the image data recorded as a slanting track on magnetic tape. Compression by DCT (discrete cosine transformation) is carried out, and two kinds of error correcting codes, outside numerals and an inner code, are used together, shuffling for one scanning line before generating outside numerals, and shuffling in 1 sector before generating an inner code after adding outside numerals are performed, and it is recorded on a recording medium. 204-block image data is recorded on one track. This block comprises a 2 words synchronous code, a 2 words ID code, an 85 words image code, 8 words inner code parity, an 85 words image code, and 8-word inner code parity. The voice data can record the numerals up to 20 bits sampled at 48 kHz, serves as the almost same composition as dynamic image data, and as shown in a figure, it is recorded on the both ends of dynamic image data by the recording medium via a gap.

[0005]

[Problem(s) to be Solved by the Invention] Thus, when video and a sound were recorded on a disk or a tape, in the disc medium represented by an optical disc, HDD etc., it could not say that seek time was quick enough, and there was a problem to which seek time becomes very long in tape media to the extent that even the concept of seeking cannot be found. In order for there to be such restrictions and to make processing of a regenerative signal easy -- real time -- and each data is sequentially multiplexed and recorded by fixed length inevitably.

[0006] On the other hand, since seek time is dramatically as high-speed as there is no concept of seeking itself, the IC memory card does not need to manage dynamic image data and voice data as fixed length data. Since a scene thru/or a soundless state with little data volume also exists in order to make data volume into the minimum, it is rational to adopt an unfixed length method as the compressing method. When an unfixed length method is adopted, if it is only the continuous reproduction from a head, trouble will not be produced by the method of managing video and a sound as another file, either, but if it performs that it is equivalent to the reproduction from the middle and rewinding to the middle, a certain hour entry is required.

[0007] Therefore, the purpose of this invention is to provide the picture/voice recording device and recording and reproducing device using the semiconductor memory which can take the synchronization with video and a sound as a recording medium using semiconductor memory, such as an IC memory card.

[0008]

[Means for Solving the Problem] In a picture and a voice recording device using semiconductor memory to which the invention according to claim 1 carries out record reproduction of image data and voice data which were inputted, an image encoding means which performs compression encoding to image data and generates variable-length coded image data. It is a picture and a voice recording device using semiconductor memory consisting of a voice coding means to perform compression encoding to voice data and to generate variable-length coded voice data, and a recording device which records coded image data and coded voice data on removable semiconductor memory.

[0009] In a picture and voice recording playback equipment using semiconductor memory to which the invention according to claim 4 carries out record reproduction of image data and voice data which were inputted, an image encoding means which performs compression encoding to image data and generates variable-length coded image data, a voice coding means to perform compression encoding to voice data and to generate variable-length coded voice data, a recording device which records coded image data and coded voice data on removable semiconductor memory, an image decoding means which reads coded image data from semiconductor memory, decodes read coded image data, and generates reproduced image data. It is an image and voice recording playback equipment using semiconductor memory consisting of a speech decoding means which reads coded voice data from semiconductor memory, decodes read coded voice data, and generates playback voice data.

[0010] It is coded, for example, by MPEG and dynamic image data is memorized by IC memory card by GOP units. Synchronizing with the GOP units, it is coded, for example, by ATRAC. Voice data is memorized by IC memory card, and hour entry management for taking a synchronization with compressed image data and compression audio data further is memorized. By this, when reproducing arbitrary positions, compression audio data in sync with compressed image data and its compressed image data can be arbitrarily read from hour entry management, and it can reproduce.

[0011]

[Embodiment of the Invention] Hereafter one example of this invention is described in detail with reference to drawings. The composition of one example of this invention is shown in drawing 1. A sound is inputted from the microphone shown by 1. The inputted sound is changed into an electrical signal and supplied to the amplifier 2 as an audio signal by the sound / electrical signal sensing element contained in the microphone 1. In the amplifier 2 an audio signal is amplified and it is restricted to a required zone. Then an audio signal is digitized by a zone more than twice the sampling frequency of being required in the A/D conversion circuit 3. It is coded by the voice compression section of voice data compression / expansion circuit 4 and this voice data is recorded on the memory 14 which consists of DRAMs etc.

[0012] The compression audio data recorded on the memory 14 is read via the card I/F circuit 15 and is recorded with IC memory card 16. This IC memory card 16 is one of the storages which can be provided with the function of a common memory card and can be written with a personal computer. As for IC memory card 16 record/reproduction is made in the format of DOS form as an example. The connector which equips with IC memory card 16 between the card I/F circuit 15 and IC memory card 16 is allotted actually and attachment and detachment of an IC memory card are enabled by the connector.

[0013] The compression audio data recorded on IC memory card 16 is read via the card I/F circuit 15. The read compression audio data is written in the record section for the sounds of the memory 14. The written-in compression audio data is elongated by real time in the voice expanding part of voice data compression / expansion circuit 4. The elongated voice data is supplied to D/A converter 5 and is analog-ized. The analog-ized audio signal is amplified with the amplifier 6 and is outputted from the loudspeaker 7 as an audio signal.

[0014] The photographic subject used as video is inputted from a lens and the converging section 8. A photographic subject is inputted into CCD image sensor 9 from a lens and the converging section 8. In CCD image sensor 9 a photographic subject is accumulated as an electric charge and changed into an electrical signal. An electrical signal is supplied to CDS circuit 10 as a dynamic image signal. CDS circuit 10 is a correlation double sampling circuit and aims at reduction of a noise. moreover -- CPU -- 22 -- a lens -- a converging section -- eight -- a mechanical system -- a driver -- 18 -- a CCD image sensor -- nine -- a CDS circuit -- ten -- and -- an A/D converter -- 11 -- an electric system -- a timing signal generation part -- (TG) -- 19 -- controlling .

[0015] The output signal of CDS circuit 10 is supplied to A/D converter 11 and is digitized with A/D converter 11 by a sampling frequency which serves as an integral multiple of the optimal sampling frequency for example the subcarrier frequency of an NTSC signal. The output signal of digitized CDS circuit 10 is supplied to the digital signal processing circuit 12. In the digital signal processing circuit 12 the usual camera signal processings such as a gamma correction and color separation is made to the supplied signal and the luminance signal Y the color-

difference signal Cband Cr are created. It is coded by the image compression part of image data compression / expansion circuit 13 and these signals are recorded on the memory 14 which is a storage cell. The compressed image data recorded on the memory 14 is read via the card I/F circuit 15 and is outputted and recorded on IC memory card 16.

[0016] The compressed image data recorded on IC memory card 16 is read via the card I/F circuit 15. The read compressed image data is written in the record section for the pictures of the memory 14. The written-in compressed image data is elongated by real time in the picture expanding part of image data compression / expansion circuit 13. The elongated dynamic image data is supplied to the digital signal processing circuit 12. The luminance signal Y and the color-difference signal Cb which were supplied to the digital signal processing circuit 12 and Cr are outputted to LCD (liquid crystal display) 17 as a video signal by which digital decoding was carried out and D/A conversion was carried out to the NTSC signal.

[0017] The operation inside the camera of CPU 22 is controlled according to external operation directions of the directions from the final controlling element 21 or a remote control (not shown). The display of a camera internal state etc. is displayed on the indicator 20 which comprises LCD, LED etc. for example.

[0018] An example of the memory area of IC memory card 16 is shown in drawing 2. This IC memory card comprises the boot area M1, the memory allocation table region M2, the directory information table region M3, the file information table M4 and the file data area M5. Boot block information is recorded on the boot area M1. To this boot block information, the version of this IC memory card, the pointer of a title, a password, a rewriting date, a rewriting mode, a partition information, a root directory and a memory allocation, the pointer of size and directory information, the pointer of size and file information, size etc. are recorded.

[0019] Memory-block information is recorded on the memory allocation table region M2. The information on each block that this memory-block information divided the record section of this IC memory card into the block of the plurality of predetermined size is recorded. In for example, the case so that the file data recorded by the user may be recorded ranging over two or more blocks. When more information which shows whether it is the block to which it was linked from the leading block, whether this block is a leading block, information on being that this block is used and this block than the number of erase times set up beforehand are eliminated, the information which considers the reliability of data and is made impossible [use] is recorded.

[0020] Whether it is the number of erase times of this block and the type of this block, for example, the block currently used for the boot area. ** is recorded [whether it is the block with which the file data is recorded / whether it is the block currently used / whether it is the block currently used / whether it is a memory area and / for the directory information table M3 and / for the file information table M4 and / by the user and] as information. As mentioned above, when there is a block linked, the address of the following block is recorded as information and as for an end ***** case, the information of an end is recorded

with this block.

[0021]The directory information used as the directory currently used for the directory information table for memory-block information is recorded on the directory information table region M3. In for example the case so that the file data recorded by the user may be recorded ranging over two or more directories. The information which shows whether it is the directory to which it was linked from the top directory whether this directory is a head When the information on being that this directory is used and the number of the files recorded in this directory are set up beforehand As the information on the file already recorded is recorded as the element 0 (file number) the element 1 (file number) ... the element n (file number) and was mentioned above When there is a directory linked the address of the next directory is recorded as information and when having ended by this directory the information of an end is recorded.

[0022]The file information considered as the file currently used for the file information table for memory-block information is recorded on the file information table region M4. For example. [whether this file is used as a file whether it is used as a directory and] It is recorded whether it is intact and in the case of whether owner ID the creation date when this file was created or the corrected modification date and this file are carried out to a read-only file or it is made the file which can be overwritten and a program it is recorded whether it prevents from performing. When the size of this file the name the number of memory block or the number of a directory and the file are recorded an attribute creator information etc. on a file are recorded. With this creator information it is recorded with what kind of application for example this file was created or with what kind of application it is usable in this file.

[0023]The file data currently used for the file data by the user for memory-block information is recorded on the file data M5. As for this file data compressed image data compression audio data etc. are recorded. As an example of these controlling methods the absolute address of IC memory card 16 is used.

[0024]In this one example in image data compression / expansion circuit 13 in order to compress dynamic image data coding / decoding of MPEG 2 are used and in order to compress voice data coding / decoding of ATRAC are used in voice data compression / expansion circuit 4. The dynamic image data and voice data which were coded are recorded on the file data M5 of IC memory card 16 mentioned above. An example of the data format is explained. IC memory card 16 is recorded in the format which records / is easy to reproduce with the operation system by the side of a personal computer in order to enable record / reproduction with a personal computer. Generally a DOS format is used. Therefore DOS file form is applied also in this example.

[0025]Here ATRAC is explained briefly. The audio signal of an analog is sampled by a 44.1-kHz sampling frequency the digitized voice data is cut down by the time window for a maximum of 11.6 ms and it decomposes into a frequency component by modification DCT operation. A frequency band is divided into three in order to avoid the preecho which is easy to produce in modification DCT operation. After

changing into a frequency axis by modification DCT data is operated on a curtailed schedule using human being's aural characteristic. It leaves the ingredient which can be heard since the noise or the small signal near the large music signal frequency component cannot be heard if auditory masking is used and the ingredient which cannot be heard and which is not required is cut. and compressed data -- 1 sound group -- (--- it bundles to every SG) and is recorded on a recording medium by a cluster unit.

[0026] This ATRAC is adopted by MD (mini disc) and is the method of compression of the voice data standardized. In this MD the mode 2 of CD-ROM is used and compression audio data is recorded on the magneto-optical disc for sound recording. The data configuration of MD recorded on the magneto-optical disc for sound recording is shown in drawing 3. In this magneto-optical disc for sound recording as shown in drawing 3 A 36 sectors are recorded as solidifying with ** (= one cluster). Three top sectors are made into a link sector region 4 sector eye is made into a sub-data sector region and the 32 remaining sectors are treated as a field for compressed data. As shown in drawing 3 Bone sector consists of 2352 bytes among those the object for data is 2332 bytes. Compression audio data treats 424 bytes as one unit and calls the unit a sound group. As shown in drawing 3 C 11 sound groups are recorded on two sectors. And if 1 sound group's compression is solved and it returns to time base information as shown in drawing 3 D it will become 512 samples of a right-and-left channel.

[0027] As shown in drawing 4 it is a case where gave MPEG 2 to dynamic image data and ATRAC is given to voice data and a video file and a voice file are considered as another file and three files of the hour entry management file for taking the synchronization with dynamic image data and voice data become 1 set of data. The video file shown in drawing 4 A comprises a header area and a field of dynamic image data. The record time screen size a frame number record bit stream ID picture achieving etc. of the dynamic image data currently recorded are recorded on the header area of the graphics file as an example. Picture achieving is a ratio of the luminance signal Y and the color-difference signal C and the sampling frequency of Cr for example it is set to 4:2:2 or 4:1:1. Dynamic image data is divided into the GOP units of MPEG in this one example.

[0028] The voice file shown in drawing 4 B comprises a header area and a field of voice data. The record time a sampling frequency a quantifying bit number record bit stream ID etc. of the voice data currently recorded are recorded on the header area of the voice file as an example. Voice data is divided per audio group (AG) who collected the clusters of ATRAC mentioned above.

[0029] And the hour entry management file shown in drawing 4 C comprises a header area and a field of voice data similarly. Data size data type ID etc. of a hour entry management file which are memorized are recorded on the header area of a hour entry management file as an example. Hour entry data is time management information and is recorded by the contents which show below the data address of the voice data which comprised a block of fixed size for example synchronized with the dynamic image data in GOP units.

[0030]The start address of GOP-N is recorded and AG number (M) which synchronizes with GOP-N is recorded. And the start address of AG-M is recorded and the offset in the AG-M is recorded.

[0031]Concretely voice data is divided into an audio group on the basis of GOP of MPEG 2 which compresses into dynamic image data. An example of GOPan audio group and a hour entry file is shown in drawing 5. As shown in drawing 5 A in the case of a video file MPEG 2 is compressed into dynamic image data and it is recorded by GOP units. Since MPEG 2 of the unfixed length compressing method is used the length of the record section of the data corresponding to each GOP differs. Thus dynamic image data is recorded per several frames (GOP units) from which each data size differs.

[0032]As shown in drawing 5 B in the case of a voice file ATRAC is compressed into voice data and it is considered as a data configuration as shown in drawing 3 B. For example the compressed audio group corresponding to 1 GOP is unfixed length. Thus voice data is also recorded by set (audio group) of GOP and a corresponding sample and this audio group's data sizes also differ.

[0033]And as shown in drawing 5 C in the case of a hour entry management file the hour entry for taking the synchronization with GOP of dynamic image data and the audio group of voice data is recorded. As for this hour entry data data is divided into GOP units. In this example a hour entry is recorded on a hour entry number (p) and as for a GOP number (n) and the audio group number (m) the start address of the voice data which synchronized with dynamic image data and its dynamic image data for every hour entry number is stored. When 1 GOP and the accompanying audio information are 1 audio group it becomes the same as that of a GOP number (n) and the audio group number (m) but 1 GOP may correspond with two or more audio groups.

[0034]Thus an example which reproduces video and a sound from the IC memory card recorded by the file from which compressed image data and compression audio data differ is explained. Drawing 6 is a block diagram showing an example of the reproduction special-purpose machine of an IC memory card. The hour entry management file by which IC memory card 32 is first reproduced by the card I/F circuit 33 is read. According to the read hour entry management file compressed image data and compression audio data are read. Compression audio data is supplied to voice data compression / expansion circuit 34 and compressed image data is supplied to image data compression / expansion circuit 38.

[0035]In image data compression / expansion circuit 36 compressed image data is elongated and the digital signal processing circuit 39 is supplied as dynamic image data. In the digital signal processing circuit 39 the supplied dynamic image data is changed into the signal of the analog-ized NTSC system and is supplied to LCD 40. A dynamic image signal is displayed in LCD 40. In voice data compression / expansion circuit 34 compression audio data is elongated and D/A converter 35 is supplied as voice data. The analog-ized audio signal takes the dynamic image signal and synchronization which are displayed on LCD 40 from the loudspeaker 37 via the amplifier 36 and is outputted. These control is performed by CPU 31.

[0036]An example of control of CPU at the time of this reproduction is explained using the flow chart of drawing 7. In Step S1a hour entry management file is read from an IC memory card and it is recorded on the memory of CPU22. An input of the position which wishes to reproduce will compute the start address (starting point) of the compressed image data which wishes to reproduce from the hour entry management file recorded on the memory in Step 2. In Step 3the start address (starting point) of the compression audio data reproduced from the hour entry management file currently recorded similarly is computed.

[0037]In step S4compressed image data is read from the start address of the computed compressed image data. The read compressed image data is sent to an elongation processing part. In Step S5compression audio data is read from the start address of the computed compression audio data. The read compression audio data is sent to an elongation processing part. When voice data has offset at this time an offset part shifts voice data and it is elongated. By this the synchronization of the dynamic image data and voice data which were reproduced is taken by GOP units.

[0038]In this one example although synchronized on the basis of dynamic image data it may synchronize on the basis of voice data.

[0039]In this one example it is explained on the assumption that management what is called by a DOS file but management by the other managing system is realizable similarly.

[0040]

[Effect of the Invention]If it depends on this invention when reproducing the video and sound of arbitrary time the hour entry file of that time can be read dynamic image data and voice data can be synchronized according to the start address of the dynamic image data and voice data which were recorded on that hour entry file and it can reproduce.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1]It is a block diagram of one example of this invention.

[Drawing 2]It is a schematic diagram of the memory of the IC memory card concerning this invention.

[Drawing 3]It is an approximate line figure used for explanation of ATRAC.

[Drawing 4]It is an approximate line figure used for explanation of the data recorded on the IC memory card concerning this invention.

[Drawing 5]It is an approximate line figure used for explanation of the data recorded on the IC memory card concerning this invention.

[Drawing 6]It is a rough block diagram for explaining an example at the time of reproduction of this invention.

[Drawing 7]It is a flow chart for explaining an example at the time of reproduction of this invention.

[Drawing 8] It is an approximate line figure used for explanation of MPEG.

[Drawing 9] It is an approximate line figure used for explanation of digital VCR.

[Description of Notations]

1 ... A microphone 26 ... An amplifier 311 ... A/D converter 4 ... Voice data
compression / expansion circuit 5 ... A D/A converter 7 ... Loudspeaker 8 ... A lensa
converging section 9 ... A CCD image sensor 10 ... CDS circuit 12 ... A digital signal
processing circuit 13 ... Image data compression/expansion circuit 14 [... LCD 18 / ...
A driver 19 / ... A timing signal generation part 20 / ... An indicator 21 / ... A final
controlling element 22 / ... CPU] ... A memory circuit 15 ... A card I/F circuit 16 ...
An IC memory card 17
